

多分辨率矢量量化的音频编解码方法及装置

技术领域

本发明涉及信号处理领域，具体地说，涉及对音频信号实现多分辨率分析和矢量量化的
5 编解码方法及装置。

背景技术

一般地，音频编码方法包括心理声学模型计算、时频域映射、量化和编码等步骤，其中
时频域映射是指将音频输入信号从时间域映射到频率域或时-频域。

时频域映射又称作变换和滤波，是音频信号编码的一个基本操作，可以提高编码效率。
10 通过此操作，时域信号包含的大部分信息都能够被转换或集中到频域或时频域系数的一个子
集中。知觉音频编码器的一个基本操作是把输入的音频信号从时间域映射到频率域或时-频
域，其基本的思路为：把信号分解为各频率带上的成分；一旦输入信号在频域上得以表达，
心理声学模型就可以用来去除感知无关信息；然后将各频带上的成份分组；最后通过合理地
分配比特数以表达各组频率参数。如果音频信号展现出较强的准周期性，这一过程可大大降
15 低数据量、提升编码效率。目前常用的时频域映射方法有：离散傅立叶变换DFT法、离散余
弦变换DCT法、镜像滤波器QMF法、伪镜像滤波器PQMF法、余弦调制滤波器CMF法、修正离散
余弦变换MDCT和离散小波（包）变换DW（P）T法等，但上述方法或者是采用一种变换/滤波
配置去压缩表达一个输入信号帧，或者是采用时域分析区间较小的滤波器组或变换压缩来表
达变化剧烈的信号，以消除前回声对解码信号的影响。而当一个输入信号帧包含不同暂态特
20 性的成份时，单一的变换配置无法满足不同信号子帧对优化压缩的基本需求；简单地采用时
域分析区间较小的滤波器组或变换来处理快变信号，则所得系数的频率分辨率较低，使得低
频部分的频率分辨率远大于人耳的临界子带带宽，严重影响了编码效率。

在音频编码过程中，当时域信号映射为时频域信号后，采用矢量量化技术可以提高编码
效率。目前在音频编码中应用矢量量化技术的音频编码方法是变换域加权交叉矢量量化
25 （Transform-domain Weighted Interleave Vector Quantization，简称TWINVQ）编码方法，
该方法在对信号进行MDCT变换后，通过交叉选择信号谱参数构造待量化的矢量，然后采用高
效率的矢量量化使较低码率的编码音频质量获得明显提高。但是，由于无法有效控制量化噪
声和人耳掩蔽的关系，TWINVQ编码方法本质上是一个感知有损的编码方法，在追求更高的主
观音频质量时，TWINVQ编码方法需要进一步的改进。同时，由于TWINVQ编码方法在组织矢量

时采用系数交织的方式，虽然可以保证矢量间统计的一致性，但对于信号能量在局部时频区域集中的现象，不能有效的利用，也限制了编码效率的进一步提高。而且，由于MDCT变换实质上是一种等带宽的滤波器组，因此，不能按照信号能量在时频平面的聚集性对信号进行分解，限制了TWINVQ编码方法的效率。

- 5 因此，如何有效利用信号的时-频域局部集聚性和矢量量化技术的高效率，是提高编码效率的一个核心问题，具体涉及两个方面：首先，需要对时频平面进行有效划分，使得信号成分同类间距离尽可能大，而类内距离尽可能小，这是解决信号的多分辨率滤波问题；其次，需要在一个有效的时频平面划分的基础上重新组织、选择和量化矢量，使得编码增益最大，这是解决信号的多分辨率矢量量化问题。

10 发明内容

本发明所要解决的技术问题在于提供一种多分辨率矢量量化的音频编解码方法及装置，可以针对不同的输入信号类型，调整时频分辨率，并有效利用信号的时频域局部集聚性进行矢量量化，提高编码效率。

- 15 本发明所述多分辨率矢量量化的音频编码方法，包括：对输入的音频信号进行自适应滤波，获得时频滤波系数，输出滤波信号；对上述滤波信号在时频平面上进行矢量划分，获得矢量组合；选择进行矢量量化的矢量；对选择的矢量进行矢量量化，并计算量化残差；量化后的码本信息作为编码器的边信息传输到音频解码器，对量化残差进行量化编码。

- 20 本发明所述多分辨率矢量量化的音频解码方法，包括：从码流中解复用得到多分辨矢量量化的边信息，获得选择点的能量以及矢量量化的位置信息；根据上述信息用逆矢量量化获得归一化的矢量，并计算归一化因子，重构出原始时频平面的量化矢量；根据位置信息将上述重构的矢量加到对应时频系数的残差上；经过多分辨率逆向滤波和频率到时间的映射，得到重构的音频信号。

- 25 本发明所述多分辨率矢量量化的音频编码器，包括时间-频率映射器、多分辨率滤波器、多分辨率矢量量化器、心理声学计算模块和量化编码器；所述时间-频率映射器接收音频输入信号，进行时间到频率域的映射，并输出到所述多分辨率滤波器；所述多分辨率滤波器用于对进行自适应滤波，输出滤波后的信号到所述心理声学计算模块和所述多分辨率矢量量化器；所述多分辨率矢量量化器用于对滤波后的信号进行矢量量化并计算量化残差，将量化后的信号作为边信息传给音频解码器，将量化残差输出到所述量化编码器；所述心理声学计算模块用于根据输入的音频信号计算心理声学模型的掩蔽阈值，并输出到所述量化编码器，用于控制量化容许的噪声；所述量化编码器用于在所述心理声学计算模块输出的容许噪声限制
- 30

下, 对所述多分辨率矢量量化器输出的残差进行量化和熵编码, 得到编码的码流信息。

本发明所述多分辨率矢量量化的音频解码器, 包括解码和逆量化器、多分辨率逆矢量量化器、多分辨率逆向滤波器和频率-时间映射器; 所述解码和逆量化器, 用于对码流解复用、熵解码和逆量化, 得到边信息及编码数据, 输出到所述多分辨率逆矢量量化器中; 所述多分辨率逆矢量量化器, 用于进行逆矢量量化过程, 重构量化的矢量, 并且将重构矢量加到时频平面上的残差系数, 输出到所述多分辨率逆向滤波器; 所述多分辨率逆向滤波器, 用于对所述多分辨率矢量量化器重构的矢量和残差系数的和信号进行逆向滤波, 并输出到所述频率-时间映射器; 所述频率-时间映射器, 用于完成信号从频率到时间的映射, 得到最终重构的音频信号。

10 本发明所述基于多分辨率矢量量化(Multiresolution Vector Quantization, 简称 MRVQ)技术的音频编解码方法及装置, 可以自适应地对音频信号进行滤波, 通过多分辨率滤波, 可以更有效的利用信号能量在局部时频区域集中的现象, 并且可以根据信号的类型, 自适应的调整时间和频率分辨率; 通过对滤波系数重新进行组织, 可以按照信号的聚集特性选择不同的组织策略, 有效的利用上述多分辨时频分析的结果; 采用矢量量化来量化这些区域, 既能
15 提高编码效率, 也能方便地控制量化的精度并进行优化。

附图说明

图 1 是本发明多分辨率矢量量化音频编码方法的流程图;

图 2 是本发明编码方法中多分辨率滤波的流程图;

图 3 是基于余弦调制滤波的信源编/解码系统的示意图;

20 图 4 是经过多分辨率滤波后能量的三种聚集模式示意图;

图 5 是多分辨率矢量量化过程的流程图;

图 6 是按照三种方式划分矢量的示意图;

图 7 是多分辨率矢量量化的一个实施例的流程图;

图 8 是区域能量/最大值的示意图;

25 图 9 是多分辨率矢量量化的另一个实施例的流程图;

图 10 是本发明多分辨率矢量量化音频编码器的结构示意图;

图 11 是音频编码器中多分辨率滤波器的结构示意图;

图 12 是音频编码器中多分辨率矢量量化器的结构示意图;

图 13 是本发明多分辨率矢量量化音频解码方法的流程图;

30 图 14 是多分辨率逆向滤波的流程图;

图 15 是本发明多分辨率矢量量化音频解码器的结构示意图;

图 16 是音频解码器中多分辨率逆矢量量化器的结构示意图;

图 17 是音频解码器中多分辨率逆向滤波器的结构示意图。

具体实施方式

5 下面根据附图及实施例进一步详细说明本发明的技术方案。

图 1 所示的流程图给出了本发明音频编码方法的总体技术方案, 输入的音频信号首先经过多分辨率的滤波, 然后对滤波系数重新进行组织, 在时频平面上进行矢量划分; 再进一步选择确定需要进行量化的矢量; 确定了矢量后, 对每个矢量进行量化, 获得相应的矢量量化码本和量化残差。矢量量化码本作为边信息发给解码器, 而量化残差则进行量化编码处理。

10 对音频信号进行多分辨率滤波的流程图如图 2 所示, 将输入的音频信号分解成帧, 对信号帧进行暂态性度量计算, 通过比较暂态性度量的值与阈值的大小来判断当前信号帧的类型是缓变信号还是快变信号。根据不同信号帧的类型选择信号帧的滤波结构, 如果是缓变信号, 则进行等带宽的余弦调制滤波, 获得时频平面的滤波系数, 输出滤波信号。如果是快变信号, 则进行等带宽的余弦调制滤波, 获得时频平面的滤波系数, 再采用小波变换对滤波系数进行
15 多分辨率分析, 调整滤波系数的时频分辨率, 最后输出滤波信号。对于快变信号, 还可以进一步地定义一系列的快变信号类型, 即存在多个阈值对快变信号进行细分, 对不同类型的快变信号, 采用不同的小波变换进行多分辨率分析, 如小波基可以是固定的, 也可以是自适应的。

如上所述, 对缓变信号和快变信号的滤波均是基于余弦调制滤波器组的技术, 余弦调制
20 滤波器组包括两种滤波形式: 传统的余弦调制滤波技术和修正离散余弦变换 MDCT 技术。基于余弦调制滤波的信源编/解码系统如图 3 所示。在编码端, 输入信号被分析滤波器组分解成 M 个子带, 将子带系数量化和熵编码。在解码端, 经熵解码和反量化后, 获得子带系数, 子带系数通过综合滤波器组滤波, 恢复音频信号。

传统的余弦调制滤波技术的冲击响应为:

$$25 \quad h_k(n) = 2p_a(n) \cos\left(\frac{\pi}{M}(k+0.5)(n-\frac{D}{2}) + \theta_k\right) \quad (F-1)$$

$$n=0,1,\dots,N_h-1$$

$$f_k(n) = 2p_s(n) \cos\left(\frac{\pi}{M}(k+0.5)(n-\frac{D}{2}) - \theta_k\right) \quad (F-2)$$

$$n=0,1,\dots,N_f-1$$

其中 $0 \leq k < M-1$, $0 \leq n < 2KM-1$, K 为大于零的整数, $\theta_k = (-1)^k \frac{\pi}{4}$ 。这里, 设 M 个子带余弦调制滤波器组的分析窗 (分析原型滤波器) $p_a(n)$ 的冲击响应长度为 N_a , 综合窗 (或称综合原型滤波器) $p_s(n)$ 的冲击响应长度为 N_s , 此时整个系统的延时 D 可限定在 $[M-1, N_s + N_a - M + 1]$ 的范围内, 系统的延时为 $D = 2sM + d (0 \leq d \leq 2M-1)$ 。

5 当分析窗和综合窗相等, 即

$$p_a(n) = p_s(n), \text{ 且 } N_a = N_s \quad (\text{F-3})$$

时, 公式 (F-1) 和 (F-2) 表示的余弦调制滤波器组为正交滤波器组, 此时矩阵 H 和 F ($[H]_{n,k} = h_k(n), [F]_{n,k} = f_k(n)$) 为正交变换矩阵。为获得线性相位滤波器组, 进一步规定对称窗

$$10 \quad p_a(2KM-1-n) = p_a(n) \quad (\text{F-4})$$

为保证正交和双正交系统的完全重构性, 窗函数需满足的条件见文献 (P.P. Vaidynathan "Multirate Systems and Filter Banks", Prentice Hall, Englewood Cliffs, NJ, 1993)。

另一种滤波形式为修正离散余弦变换 MDCT, 也被称为 TDAC (Time Domain Aliasing Cancellation) 余弦调制滤波器组, 其冲击响应为:

$$15 \quad h_k(n) = p_a(n) \sqrt{\frac{2}{M}} \cos\left(\frac{\pi}{M}(k+0.5)(n + \frac{M+1}{2})\right) \quad (\text{F-5})$$

$$f_k(n) = p_s(n) \sqrt{\frac{2}{M}} \cos\left(\frac{\pi}{M}(k+0.5)(n + \frac{M+1}{2})\right) \quad (\text{F-6})$$

其中 $0 \leq k < M-1$, $0 \leq n < 2KM-1$, K 为大于零的整数。其中, $p_a(n)$ 和 $p_s(n)$ 分别为分析窗 (或分析原型滤波器) 和综合窗 (或综合原型滤波器)。

同样的, 当分析窗和综合窗相等, 即

$$20 \quad p_a(n) = p_s(n) \quad (\text{F-7})$$

时, 公式 (F-5) 和 (F-6) 表示的余弦调制滤波器组为正交滤波器组, 此时矩阵 H 和 F ($[H]_{n,k} = h_k(n), [F]_{n,k} = f_k(n)$) 为正交变换矩阵。为获得线性相位滤波器组, 进一步规定对称窗

$$p_a(2KM-1-n) = p_a(n) \quad (\text{F-8})$$

25 则为满足完全重构, 由此可知, 分析窗和综合窗需满足

$$\sum_{m=0}^{2K-1-2s} p_a(mM+n) p_a((m+2s)M+n) = \delta(s) \quad (\text{F-9})$$

其中 $s=0, \dots, K-1$, $n=0, \dots, \frac{M}{2}-1$ 。

放宽公式 (F-7) 的约束条件, 即取消分析窗和综合窗相等的限制, 则余弦调制滤波器组为双正交调制滤波器组。

- 5 时域分析已经证明, 根据公式 (F-5) 和 (F-6) 获得的双正交调制滤波器组依然满足完全重构性能, 只要

$$\sum_{m=0}^{2K-1-2s} p_s(mM+n) p_a((m+2s)M+n) = \delta(s) \quad (\text{F-10})$$

$$\sum_{m=0}^{2K-1-2s} (-1)^m p_s(mM+n) p_a((m+2s)M+(M-n-1)) = 0 \quad (\text{F-11})$$

其中 $s=0, \dots, K-1$, $n=0, \dots, M-1$ 。

- 10 根据上述分析, 余弦调制滤波器组 (包括MDCT) 的分析窗和综合窗均可以采用任意满足滤波器组完全重构条件的窗形式, 如在音频编码中常用的SINE和KBD窗。

- 另外, 余弦调制滤波器组滤波可以采用快速傅立叶变换来提高计算效率, 可参考文献 “A New Algorithm for the Implementation of Filter Banks based on ‘Time Domain Aliasing Cancellation’” (P. Duhamel, Y. Mahieux和J. P. Petit, Proc. ICASSP, 1991年五月, 2209-2212 页)。

同样, 小波变换技术也是信号处理领域众所周知的技术, 可以参考 “子波变换理论及其在信号处理中的应用” (陈逢时, 国防工业出版社, 1998) 关于小波变换技术的详细论述。

- 20 经过多分辨率分析滤波后的信号在时间-频率平面上具有重新分配、聚集信号能量的性质, 如图 4 所示。对时域平稳的信号, 如正弦信号, 在时频平面上, 其能量会沿时间方向聚集在一个频率带上, 如图 4 的 a 所示; 对时域快变信号, 尤其是音频编码中预回声现象明显的快变信号, 如响板信号, 其能量沿频率方向分布为主, 即大部分的能量值聚集在少数几个时间点上, 如图 4 的 b 所示; 而对于时域噪声信号, 其频谱分布在较宽的范围, 因此能量聚集方式具有多种模式, 既有沿时间方向的分布, 也有沿频率方向的分布, 还有按区域式的分布, 如图 4 的 c 所示。

- 25 在时间-频率的多分辨率分布中, 低频部分的频率分辨率高, 中高频部分的频率分辨率较低。由于引起预回声现象的成分主要是中高频部分, 如果能改善这些成分的编码质量, 就能够有效抑制预回声, 多分辨率矢量量化的一个重要出发点, 就是针对这些重要的滤波系数,

优化量化引入的误差。因此,对这些系数采用高效的编码策略特别重要。根据多分辨率滤波后得到的信号滤波系数的时间-频率分布,可以有效的把重要的滤波系数进行重组和分类。由上述分析可知,经过多分辨率滤波后的信号的能量分布呈现较强的规律,引入矢量量化,可以有效的利用这种特点进行系数的组合。通过对采用特定方式的矢量组织,把时间-频率平面上的区域组织为一维矢量的矩阵形式。然后,对此矢量矩阵的全部或部分矩阵元素实行矢量量化,量化后的信息作为编码器的边信息传输到解码器,而量化残差和未量化的系数则一起构成一个残差系统,进行量化编码。

图 5 详细描述了音频信号在经过多分辨率滤波后,进行多分辨率矢量量化的过程,多分辨率矢量量化的过程包括矢量划分、选择矢量和矢量量化三个子过程。

10 对时频平面可按照时间方向、频率方向和时频区域三种方式进行矢量划分,音调性较强的信号适用于按时间方向组织矢量,时域具有快变特性的信号则适合于按频率方向组织矢量,而比较复杂的音频信号则适合按时频区域组织矢量。假设信号的频率系数长度是 N ,经过多分辨率滤波后,在时频平面上时间方向的分辨率为 L ,频率方向的分辨率为 K ,且 $K*L=N$ 。当进行矢量划分时,首先确定矢量维数 D 的大小,由此可得到划分后的矢量的个数为 N/D 。

15 当按照时间方向进行矢量划分时,保持频率方向的分辨率 K 不变,对时间进行划分;当按照频率方向进行矢量划分时,保持时间方向的分辨率 L 不变,对频率进行划分;当按照时频区域进行矢量划分时,其时间和频率方向划分的个数可任意,只要满足最终划分的矢量个数为 N/D 即可。图 6 示出了按照时间、频率和时频区域划分矢量的实施例。假设频率系数长 $N=1024$,经过多分辨率滤波后,时频平面被划分为 $K*L=64*16$ 形式, $K=64$ 为频率方向的分辨率, $L=16$ 为时间方向的分辨率。假设矢量的维数 $D=8$,可以对该时频平面按照不同的方式组合和提取矢量,如图 6-a、图 6-b 和图 6-c 所示。在图 6-a 中,矢量按频率方向被划分为 $8*16$ 个 8 维矢量,简称为 I 型矢量组织。图 6-b 是按照时间方向划分矢量的结果,共有 $64*2$ 个 8 维矢量,简称为 II 型矢量组织。图 6-c 是按照时频区域组织矢量的结果,共有 $16*8$ 个 8 维矢量,简称为 III 型矢量组织。这样按不同的划分方法皆可获得 128 个 8 维矢量。可将 I 型组织得到的矢量集合记为 $\{v_i\}$,II 型组织得到的矢量集合记为 $\{v_j\}$,III 型组织得到的矢量集合记为 $\{v_{i,j}\}$ 。

在进行矢量划分后,然后确定需要量化哪些矢量,对矢量进行选择,可以采用两种选择方式。

30 第一种方式是选择整个时间-频率平面上的全部矢量进行量化,全部矢量是指按照某一种划分所得的覆盖全部时频格点的矢量,如可以是 I 型矢量组织得到的全部矢量,或是 II

型矢量组织得到的全部矢量,或是 III 型矢量组织得到的全部矢量,只要选择其中一组的全部矢量即可。至于选择哪一组的矢量集合,则通过量化增益来确定,量化增益是指量化前的能量与量化误差能量的比。对于上述矢量组织,选择增益值大的矢量组织的矢量。

第二种方式是选择最重要的矢量进行量化,最重要的矢量既可以包括频率方向的矢量,也可以包括时间方向的矢量或时频区域的矢量。对于只选择部分矢量进行矢量量化的情况,在边信息中除了包括矢量的量化索引外,还需包括这些矢量的序号。具体选择矢量的方法在下面的内容中进行介绍。

确定了量化的矢量后,则进行矢量量化的处理。不论是选择全部矢量进行量化,还是只选择重要矢量进行量化,其基本单元都是对单个矢量的量化。对单个 D 维矢量,考虑到动态范围和码本大小之间的折衷,需要在量化前对矢量进行归一化处理,得到一个归一化因子,归一化因子是反映不同矢量的能量动态范围的值,是变化的量。经过归一化处理后的矢量再进行量化,包含码本索引号的量化和归一化因子的量化,考虑到码率和编码增益的限制,对归一化因子的量化所占的比特数在满足精度的条件下越少越好。在本发明中,可以采用曲线和曲面拟合、多分辨率分解和预测等方法计算多分辨率时间-频率系数包络,获得归一化因子。

图 7 和图 9 分别给出了多分辨率矢量量化过程的两个具体实施例的流程图。图 7 所示实施例根据能量和矢量内分量的方差对矢量进行选择,并采用泰勒展式描述多分辨率时间-频率系数包络,获得归一化因子,再进行量化,以实现多分辨率矢量量化。图 9 所示实施例是根据编码增益来选择矢量,并采用样条曲线拟合计算多分辨率时间-频率系数包络,获得归一化因子,再进行量化,以实现多分辨率矢量量化。下面分别介绍这两个实施例。

在图 7 中,首先分别按照频率方向、时间方向和时频区域进行矢量组织,若频率系数 $N=1024$ 时,时间-频率多分辨率滤波产生 64×16 的格点,当矢量维数取 8 时,那么按频率划分可以得到 8×16 矩阵形式的矢量,按时间划分可以得到 64×2 矩阵形式的矢量,按时频区域可得到 16×8 矩阵形式的矢量。

如果不对全部矢量都量化,那么需要按照重要性来选择矢量。在本实施例中,选择矢量的依据是矢量的能量和矢量内各分量的方差,在计算方差时,矢量组成元素需要取绝对值,以排除数值符号的影响。设集合 $V = \{v_i\} \cup \{v_j\} \cup \{v_{i-j}\}$, 则选择矢量的过程具体如下:首先,计算集合 V 中的每个矢量的能量 $E v_i = |v_i|^2$, 同时计算每个矢量的 $d E v_i$, $d E v_i$ 表示第 i 个矢量的各分量方差。然后将集合 V 中的元素按能量从大到小进行排序,再将上述排序后的元素按照方差从小到大进行再排序。根据信号总能量和当前选择的矢量总能量之比确定需选择的矢量

个数M, 典型的值可取 3-50 内的整数。然后选择前 M 个矢量进行矢量量化, 若同时包含有 I 型矢量组织、II 型矢量组织和 III 型矢量组织的同一区域的矢量, 则按方差的排序进行取舍。通过上述步骤, 选择出待量化的 M 个矢量。

在选择了 M 个矢量后, 利用泰勒 Taylor 近似公式, 分别用不同的失真度量准则, 完成对各阶差分的量化搜索过程。为了更有效的量化, 需要对矢量进行两次归一化处理, 第一次归一化时采用全局最大绝对值, 第二次归一化时, 通过有限多点对信号包络进行估计, 然后用估计值对对应位置矢量进行第二次归一化, 经过两次归一化后, 矢量变化的动态范围得到有效的控制。信号包络的估计方法通过泰勒展式实现, 将在后面详细叙述。

矢量量化按以下步骤进行: 首先确定 Taylor 近似计算公式中的参数, 以便使用泰勒公式来表示整个时频平面上任意矢量的能量近似值, 并且计算出其中的最大能量或最大绝对值; 然后, 对选择出来的矢量进行第一次归一化处理; 接着通过 Taylor 公式计算待矢量量化的矢量的能量近似值, 进行第二次归一化处理; 最后对归一化后的矢量按最小失真进行量化, 并计算量化残差。下面对上述步骤进行详细地描述。在时间-频率平面上, 每个时频格点上的系数对应一个确定的能量值。定义时频格点的系数能量为该系数的平方或其绝对值; 定义矢量的能量为组成该矢量的所有时频格点上系数能量的和或者这些系数值中最大的绝对值; 定义时频平面区域的能量为组成该区域的所有时频格点上系数能量的和或者这些系数值中最大的绝对值。因此为了得到矢量的能量, 需要对矢量所包含的所有时频格点系数计算能量和或者绝对值最大的值。因此, 对整个时间-频率平面, 可以采用图 6-a、6-b 和/或 6-c 的划分方式, 对划分后的区域进行编号 (1、2、...N)。如果采用按频率方向划分, 则每个区域就对应一个频率方向的矢量, 计算每个区域的能量或绝对值最大的值, 构造出一元函数 $Y=f(X)$, 其中 X 表示区域序号, 其取值为 [1, N] 上的整数, Y 表示对应 X 的区域的能量或绝对值最大的值, 而点 (X_i, Y_i) , i 取值为 [1, N] 上的整数, 也被称为引导点。根据泰勒公式有:

$$f(x_0 + \Delta) = f(x_0) + f^{(1)}(x_0)\Delta + \frac{1}{2!}f^{(2)}(x_0)\Delta^2 + \frac{1}{3!}f^{(3)}(\xi)\Delta^3 \quad (1)$$

一元函数 $Y=f(X)$ 的 M 个值构成了一个离散序列 $\{y_1, y_2, y_3, y_4, \dots, y_M\}$, 该序列的一阶、二阶和三阶差分都可以用回归方法求得, 即由 Y 可得到 DY、 D^2Y 以及 D^3Y 。

图 8 所示的是用泰勒展式近似表示函数 $Y=f(X)$ 的示意图, 圆点表示从全部 N 个区域中选择出来的待量化编码的区域, 这里的 N 是指整个时频平面划分得到的矢量数。具体获得归一化因子的过程如下: 根据信号总能量确定一个全局的增益因子 Global-Gain, 对其用对数模型量化编码。然后用该增益因子 Global-Gain 对矢量进行归一化, 再根据泰勒公式 (1) 计算出当前矢量位置上的局部归一化因子 Local-Gain, 并且对当前矢量再次进行归一化处理。

于是当前矢量的总体归一化因子 Gain 由上述两个归一化因子的乘积给出:

$$\text{Gain} = \text{Global_Gain} * \text{Local_Gain} \quad (2)$$

其中, Local_Gain 在编码器端不需要量化。在解码器端, 根据泰勒公式 (1) 用相同的过程可以求出局部归一化因子 Local_Gain。用 Global_Gain 与重构的归一化矢量相乘, 即可得到当前矢量的重构值。因此, 在编码器端需要编码的边信息就是图 8 中选择的圆点处的函数值、以及它们的一阶、二阶差分值, 本发明采用矢量量化来对它们进行编码。

矢量量化的过程描述如下: 预先选择的 M 个区域的函数值 $f(x)$ 构成 M 维矢量 y , 已知该矢量对应的一阶、二阶差分, 分别用 dy 和 d^2y 表示, 对这三个矢量分别进行量化。在编码器端, 用码本训练算法已经得到了对应三个矢量的码本, 量化过程就是搜索最佳匹配矢量的过程。矢量 y 对应泰勒公式的零阶近似表示, 在码本搜索时的失真度量用欧氏距离。对一阶差分 dy 的量化, 对应于泰勒公式的一阶近似:

$$f(x_0 + \Delta) = f(x_0) + f^{(1)}(x_0)\Delta \quad (3)$$

因此, 一阶差分的量化首先根据欧式距离, 搜索对应码本中失真最小的少量码字, 再在当前矢量 x_0 的小邻域中, 对邻域中的每一个区域用公式 (3) 计算量化失真, 最后用总的失真和作为失真度量, 即:

$$D = \sum_{k=-M}^{+M} (f(x + \Delta_k) - \hat{f}(x + \Delta_k))^2 \quad (4)$$

其中 $f(x + \Delta_k)$ 表示量化前的真值, $\hat{f}(x + \Delta_k)$ 表示用泰勒公式求出的近似值, M 表示邻域的范围。对二阶差分 d^2y 的量化可用类似的过程进行。通过上述过程最终可以得到三个量化后的码字索引, 作为边信息传输到解码器。而量化残差则进行量化编码处理。

上述方法可以很容易扩展到二维时频曲面的情况。

图 9 为多分辨率矢量量化过程的另一个具体实施例。首先分别按照频率方向、时间方向和区域进行矢量组织, 如果不对全部矢量进行量化, 则计算每个矢量的编码增益, 选择编码增益最大的前 M 个矢量进行矢量量化, M 值的确定方法是: 对矢量按照能量从大到小排序后, 占总能量百分比超过一个经验阈值 (如 50%-90%) 的矢量的数目就是 M。为了更有效的量化, 也需要对矢量进行两次归一化, 第一次采用全局最大绝对值, 第二次采用样条拟合计算矢量内归一化值, 经过两次归一化后, 矢量变化的动态范围得到有效的控制。

与图 7 所示的实施例相同, 首先对整个时间-频率平面重新进行划分并编号 (1, 2, ..., N), 计算每个区域的能量或绝对值最大的值, 构造一元函数 $Y=f(X)$, 其中 X 表示区域编号, 其取值为 [1, N] 上的整数, Y 是对应 X 的区域的能量或绝对值最大的值。根据 B 样条曲线拟合

的公式有:

第 i 个子区间上的常数 (0 次) B 样条函数为:

$$N_{i,0}(x) = \begin{cases} 1 & x_i \leq x \leq x_{i+1} \\ 0 & \text{其它} \end{cases} \quad (5)$$

5 0, 其它。

在区间 $[x_i, x_{i+m+1}]$ 上的第 m 次 B 样条函数定义为:

$$N_{i,m}(x) = \frac{(x - x_i)}{(x_{i+m} - x_i)} N_{i,m-1}(x) + \frac{(x_{i+m+1} - x)}{(x_{i+m+1} - x_{i+1})} N_{i+1,m-1}(x) \quad (6)$$

10 那么, 采用 B 样条基函数作为基底, 可以将任何样条表示为:

$$f(x) = \sum_k \beta_i N_{i,m}(x) \quad (7)$$

这样根据公式 (5)、(6) 和 (7) 可以计算给定 x 点样条的函数值, 这些用于插值的点也被称为引导点。

图 8 同样可以作为经样条曲线拟合获得的函数 $Y=f(X)$ 的示意图, 圆点表示从全部 N 个区域中选择出来的待编码的区域, 这里的 N 是整个时频平面划分得到的矢量数。具体的矢量量化过程如下: 在编码器端, 对待量化的矢量, 根据信号总能量确定一个全局的增益因子 Global-Gain, 对其用对数模型量化编码; 然后用该增益因子 Global-Gain 对矢量进行归一化, 根据拟合公式 (7) 计算当前矢量位置上的局部归一化因子 Local-Gain 并且再次对当前矢量进行归一化处理, 于是当前矢量的总体归一化因子 Gain 是上述两个因子的乘积:

$$20 \quad \text{Gain} = \text{Global-Gain} * \text{Local-Gain} \quad (8)$$

其中, Local-Gain 在编码器端并不需要量化。同样的, 在解码器端可以根据拟合公式 (7) 用相同的过程求出 Local-Gain。用总增益与重构的归一化矢量相乘, 即可得到当前矢量的重构值。因此, 在采用样条曲线拟合方法时, 编码器端需要编码的边信息就是图 8 中所选择的圆点处的函数值, 本发明采用矢量量化对它们进行编码。

25 矢量量化的过程描述如下: 预先选择 M 个区域的函数值 $f(x)$ 构成 M 维的矢量 y , 矢量 y 可以进一步分解成若干分矢量, 以控制矢量的大小, 提高矢量量化的精度, 这些矢量被称为选择点矢量。然后, 对矢量 y 分别进行量化。在编码器端, 用码本训练算法可以得到了对应的矢量码本。量化过程就是搜索最佳匹配矢量的过程, 搜索得到的码字索引作为边信息传送到解码器。量化误差则继续进行下一步的量化编码处理。

30 以上方法可以很容易扩展到二维时频曲面的情况。

如图 10 所示的音频编码器, 包括时间-频率映射器、多分辨率滤波器、多分辨率矢量量化器、心理声学计算模块和量化编码器。待编码的输入音频信号分为两路, 一路经时间-频率映射器后进入多分辨率滤波器, 进行多分辨率分析, 其分析结果作为矢量量化的输入和用于调整心理声学计算模块的计算; 另一路进入心理声学计算模块, 估计当前信号的心理声学掩蔽阈值, 用于控制量化编码器的感知不相关成分; 多分辨率矢量量化器根据多分辨率滤波器的输出, 对时频平面的系数划分成矢量并进行矢量量化, 量化残差由量化编码器进行量化和熵编码。

图 11 是图 10 所示音频编码器中多分辨率滤波器的结构示意图。多分辨率滤波器包括暂态性度量计算模块、多个等带宽余弦调制滤波器、多个多分辨率分析模块和时频滤波系数组织模块; 其中多分辨率分析模块的个数比等带宽余弦调制滤波器的个数少一个。其工作原理如下: 输入音频信号经过暂态性度量计算模块的分析, 分为缓变信号和快变信号, 快变信号可进一步细分为类 I 型快变信号, 类 II 型快变信号。对于缓变信号, 输入到等带宽余弦调制滤波器中进行滤波, 获得所需的时-频滤波系数; 对于各类快变信号, 则均先经过等带宽余弦调制滤波器进行滤波, 然后再进入多分辨率分析模块对滤波系数进行小波变换, 调整系数的时频分辨率, 最后通过时频滤波系数组织模块输出滤波后的信号。

多分辨率矢量量化器的结构如图 12 所示, 包括矢量组织模块、矢量选择模块、全局归一化模块、局部归一化模块和量化模块。多分辨率滤波器输出的时频平面系数经过矢量组织模块, 根据不同的划分策略, 组织成矢量的形式, 然后在矢量选择模块根据能量的大小等因素选择出待量化的矢量, 输出到全局归一化模块。在全局归一化该模块中, 通过全局归一化因子对所有的矢量进行第一次全局归一化处理, 然后在局部归一化模块中计算出每个矢量的局部归一化因子, 并进行第二次局部归一化处理, 输出到量化模块。在量化模块中, 对经过两次归一化后的矢量进行量化, 并计算出量化后的残差, 作为多分辨率矢量量化器的输出。

本发明还提供了多分辨率矢量量化的音频解码方法, 如图 13 所示, 首先对收到的码流进行解复用、熵解码和逆量化, 得到量化的全局归一化因子以及选择点的量化索引。根据索引从码本中计算出各个选择点的能量及各阶差分值, 从码流中得到时频平面上矢量量化的位置信息, 再根据泰勒公式或样条曲线拟合公式, 获得对应位置上的二次归一化因子。再根据矢量量化索引得到归一化的矢量, 并与上述两个归一化因子相乘, 就重构了时频平面上量化的矢量。将重构后的矢量和解码逆量化后的时频平面对应位置的系数相加, 进行多分辨率逆向滤波和频率到时间的映射, 完成解码, 得到重构的音频信号。

图 14 介绍了解码方法中的多分辨率逆向滤波的过程。首先对重构矢量的时频系数进行

时频组织, 根据解码得到的信号类型进行如下滤波操作: 如果是缓变信号, 则进行等带宽余弦调制滤波, 获得时域的脉冲编码调制 PCM 输出; 如果是快变信号, 则进行多分辨率综合, 再进行等带宽余弦调制滤波, 获得时域的 PCM 输出。对于快变信号, 也可以进一步细分为多种类型, 不同类型的快变信号进行多分辨率综合的方法也不同。

5 相应的音频解码器如图 15 所示, 具体包括解码和逆量化器、多分辨率逆矢量量化器、多分辨率逆向滤波器以及频率-时间映射器。解码和逆量化器对收到的码流进行解复用, 并进行熵解码和逆量化, 获得多分辨矢量量化的边信息, 输出到多分辨率逆矢量量化器中。多分辨率逆矢量量化器根据逆量化结果和边信息, 重构量化矢量, 并恢复时频平面的值; 多分辨率逆向滤波器对多分辨率逆矢量量化器重构的矢量进行逆向滤波, 并由频率-时间映射器
10 完成频率到时间的映射, 得到最终重构的音频信号。

上述多分辨率逆矢量量化器的结构如图 16 所示, 包括解复用模块、逆量化模块、归一化矢量计算模块、矢量重构模块和加法模块。首先解复用模块对接收到的码流进行解复用, 获得归一化因子和选择点的量化索引。然后在逆量化模块中根据量化索引获得能量包络, 根据解复用结果获得矢量量化位置信息, 并根据归一化因子和量化索引, 逆量化获得引导点和
15 选择点矢量, 计算出二次归一化因子, 输出到归一化矢量计算模块。在归一化矢量计算模块中, 对选择点矢量进行逆二次归一化, 获得归一化矢量, 输出到矢量重构模块中, 再根据能量包络对归一化矢量进行逆一次归一化, 获得重构矢量。重构矢量和对应时频平面的反量化残差在加法模块中相加, 得到逆量化的时频系数, 作为多分辨率逆向滤波器的输入。

多分辨率逆向滤波器的结构如图 17 所示, 包括时频系数组织模块、多个多分辨率综合
20 模块以及多个等带宽余弦调制滤波器, 其中多分辨率综合模块的个数比等带宽余弦调制滤波器的个数少 1。重构的矢量经过时频系数组织模块后, 分为缓变信号和快变信号, 快变信号还可以进一步的细分为多种类型, 如 I、II...K。对于缓变信号, 则输出到等带宽的余弦调制滤波器进行滤波, 获得时域 PCM 输出。对于不同的快变信号类型, 则输出到不同的多分辨率综合模块进行综合, 然后输出到等带宽的余弦调制滤波器中滤波, 获得时域 PCM 输出。

25 最后所应说明的是, 以上实施例仅用以说明本发明的技术方案而非限制, 尽管参照较佳实施例对本发明进行了详细说明, 本领域的普通技术人员应当理解, 可以对本发明的技术方案进行修改或者等同替换, 而不脱离本发明技术方案的精神和范围, 其均应涵盖在本发明的权利要求范围当中。

权利要求书

- 1、 一种多分辨率矢量量化的音频编码方法，其特征在于，包括：对输入的音频信号进行自适应滤波，获得时频滤波系数，输出滤波信号；对上述滤波信号在时频平面上进行矢量划分，
5 获得矢量组合；选择进行矢量量化的矢量；对选择的矢量进行矢量量化，并计算量化残差；量化后的码本信息作为编码器的边信息传输到音频解码器，对量化残差进行量化编码。
- 2、 根据权利要求 1 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述对音频信号进行自适应滤波的步骤进一步包括：将输入的音频信号分解成帧，计算信号帧的暂态性度量；通过比较暂态性度量的值与阈值的大小来判断当前信号帧的类型是缓变信号还是快变信号；如果是缓变信号，则进行等带宽的余弦调制滤波，获得时频平面的滤波系数，输出滤波信号；如果是快变信号，则进行等带宽的余弦调制滤波，获得时频平面的滤波系数，再采用小波变换对滤波系数进行多分辨率分析，调整滤波系数的时频分辨率，最后输出滤波信号。
10
- 3、 根据权利要求 2 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述余弦调制滤波可采用传统的余弦调制滤波或修正离散余弦变换滤波。
15
- 4、 根据权利要求 3 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述余弦调制滤波还包括进行快速傅立叶变换。
20
- 5、 根据权利要求 2 所述的多分辨率矢量量化的音频编码方法，其特征在于，如果是快变信号，则还包括：将快变信号进一步细分为多种快变信号类型，对于不同的快变信号类型，分别进行滤波和多分辨率分析。
- 6、 根据权利要求 5 所述的多分辨率矢量量化的音频编码方法，其特征在于，对不同类型的快变信号，所述进行多分辨率分析的小波变换的小波基是固定的或是自适应的。
25
- 7、 根据权利要求 1 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述对滤波信号在时频平面上进行矢量划分包括按照时间方向、频率方向和时频区域三种方式进行矢量划分；
30

所述按时间方向划分进一步包括保持频率方向的分辨率不变，对时间进行划分，使得划分后的矢量个数为 N/D ，得到 I 型矢量组织，其中 N 表示音频信号的频率系数的长度， D 表示矢量的维数；

5 所述按频率方向划分进一步包括保持时间方向的分辨率不变，对频率进行划分，使得划分后的矢量个数为 N/D ，得到 II 型矢量组织，其中 N 表示音频信号的频率系数的长度， D 表示矢量的维数；

所述按时频区域划分进一步包括对时频平面的时间和频率进行划分，使得划分后的矢量个数为 N/D ，得到 III 型矢量组织，其中 N 表示音频信号的频率系数的长度， D 表示矢量的维数。

10

8、根据权利要求 1 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述选择进行矢量量化的矢量的步骤进一步包括：判断是否需要有时频平面的全部矢量进行量化，如果是，则分别计算 I 型矢量组织、II 型矢量组织和 III 型矢量组织的量化增益，选择量化增益值大的矢量组织的矢量作为量化的矢量；如果否，则选择 M 个待量化的矢量，并对所选的矢量的
15 序号进行编码。

9、根据权利要求 8 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述选择 M 个待量化的矢量的步骤可以进一步包括：将 I 型矢量组织、II 型矢量组织和 III 型矢量组织的矢量组成一个矢量集合；计算上述矢量集合中每个矢量的能量即系数的平方，同时计算每个矢量的各分量方差；将矢量集合中的矢量按能量从大到小进行排序；将上述排序后的矢量按照方差从小到大进行再排序；根据信号总能量和当前选择的矢量总能量之比确定需选择的
20 矢量个数 M ，选择前 M 个矢量作为矢量量化的矢量；若同时包含有 I 型矢量组织、II 型矢量组织和 III 型矢量组织的同一区域的矢量，则按方差的排序进行取舍。

25 10、根据权利要求 8 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述选择 M 个待量化的矢量的步骤可以进一步包括：将 I 型矢量组织、II 型矢量组织和 III 型矢量组织的矢量组成一个矢量集合；计算矢量集合中每个矢量的能量和编码增益；选择编码增益最大的前 M 个矢量，使得所选 M 个矢量的能量与总能量的百分比超过 50%。

30 11、根据权利要求 9 或 10 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述 M

的值可以是 3 到 50 之间的任一整数。

12、根据权利要求 1 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述对选择的矢量进行矢量量化的步骤进一步包括：计算时间-频率平面每个区域的能量值或绝对值最大值；确定全局归一化因子；对选择的矢量进行归一化处理；计算矢量的局部归一化因子，并进行第二次归一化处理；对归一化后的矢量进行量化，并计算量化残差。

13、根据权利要求 12 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述对选择的矢量进行矢量量化的步骤进一步包括：计算时间-频率平面每个区域的能量值或绝对值最大值；构造一元函数 $Y=f(X)$ ，其中 X 表示区域的序号， Y 表示对应 X 的区域的能量或绝对值最大值；根据信号总能量确定一个全局增益因子，对其用对数模型进行量化编码；用该全局增益因子对选择的矢量进行归一化处理；根据泰勒公式计算当前矢量位置上的局部归一化因子，并对当前矢量再次进行归一化处理；获得当前矢量的总体归一化因子是上述两个归一化因子的乘积；将选择的 M 个区域的函数值构成 M 维矢量；计算该矢量对应的一阶、二阶差分；通过码本训练算法获得对应上述三个矢量的码本，并对上述三个矢量进行量化；所述矢量的量化对应泰勒公式的零阶近似表示，码本搜索时的失真度量采用欧氏距离；一阶差分矢量的量化对应于泰勒公式的一阶近似，根据欧式距离，搜索对应码本中失真最小的少量码字，再在当前矢量 x_0 的小邻域中，对邻域中的每一个区域计算量化失真，最后总的失真和作为失真度量；二阶差分矢量的量化与一阶差分矢量的量化类似。

14、根据权利要求 12 所述的多分辨率矢量量化的音频编码方法，其特征在于，所述对选择的矢量进行矢量量化的步骤进一步包括：计算时间-频率平面每个区域的能量值或绝对值最大值；构造一元函数 $Y=f(X)$ ，其中 X 表示区域的序号， Y 表示对应 X 的区域的能量或绝对值最大值；根据信号总能量确定一个全局增益因子，对其用对数模型进行量化编码；用该全局增益因子对选择的矢量进行归一化处理；根据样条曲线拟合公式计算当前矢量位置上的局部归一化因子，并对当前矢量再次进行归一化处理；将选择的 M 个区域的函数值构成 M 维的矢量，所述矢量可以进一步分解成若干分矢量，称为选择点矢量；对上述矢量分别进行量化。

15、一种多分辨率矢量量化的音频解码方法，其特征在于，包括以下步骤：从码流中解复用得到多分辨矢量量化的边信息，获得选择点的能量以及矢量量化的位置信息；根据上述信息

用逆矢量量化获得归一化的矢量，并计算归一化因子，重构出原始时频平面的量化矢量；根据位置信息将上述重构的矢量加到对应时频系数的残差上；经过多分辨率逆向滤波和频率到时间的映射，得到重构的音频信号。

- 5 16、根据权利要求 15 所述的多分辨率矢量量化的音频解码方法，其特征在于，所述重构原始时频平面的量化矢量步骤进一步包括：根据边信息从码本中计算出各个选择点的能量及各阶差分值；从码流中得到时频平面上矢量量化的位置信息和全局归一化因子；根据编码过程中计算二次归一化因子的公式，获得对应位置上的二次归一化因子；根据矢量化索引获得归一化的矢量，并与上述两个归一化因子相乘，重构时频平面上量化的矢量。

- 10 17、根据权利要求 15 所述的多分辨率矢量量化的音频解码方法，其特征在于，所述多分辨率逆向滤波的步骤进一步包括：对重构矢量的时频系数进行时频组织，根据解码得到的信号类型进行如下滤波操作：如果是缓变信号，则进行等带宽余弦调制滤波，获得时域的脉冲编码调制输出；如果是快变信号，则进行多分辨率综合，再进行等带宽余弦调制滤波，获得时域的脉冲编码调制输出。

- 18、根据权利要求 17 所述的多分辨率矢量量化的音频解码方法，其特征在于，所述快变信号可以进一步分为多种快变信号类型，对不同的快变信号类型，分别进行多分辨率综合和滤波。

- 20 19、一种多分辨率矢量量化的音频编码器，其特征在于，包括时间-频率映射器、多分辨率滤波器、多分辨率矢量量化器、心理声学计算模块和量化编码器；

所述时间-频率映射器接收音频输入信号，进行时间到频率域的映射，并输出到所述多分辨率滤波器；

- 25 所述多分辨率滤波器用于对信号进行自适应滤波，并输出滤波后的信号到所述心理声学计算模块和所述多分辨率矢量量化器；

所述多分辨率矢量量化器用于对滤波后的信号进行矢量量化并计算量化残差，将量化后的信号作为边信息传给音频解码器，将量化残差输出到所述量化编码器；

- 30 所述心理声学计算模块用于根据输入的音频信号计算心理声学模型的掩蔽阈值，并输出到所述量化编码器，以控制量化容许的噪声；

所述量化编码器用于在所述心理声学计算模块输出的容许噪声限制下,对所述多分辨率矢量量化器输出的残差进行量化和熵编码,得到编码的码流信息。

20、根据权利要求 19 所述的多分辨率矢量量化的音频编码器,其特征在于,所述多分辨率滤波器包括暂态性度量计算模块、M 个等带宽余弦调制滤波器、N 个多分辨率分析模块和时频滤波系数组织模块,且满足 $M=N+1$;

所述暂态性度量计算模块,用于计算音频输入信号帧的暂态性度量,以确定所述信号帧的类型;

所述等带宽余弦调制滤波器,用于对信号进行滤波,获得滤波系数;如果是缓变信号,将滤波系数输出到所述时频滤波系数组织模块;如果是快变信号,则将滤波系数输出到所述多分辨率分析模块;

所述多分辨率分析模块,用于对快变信号的滤波系数进行小波变换,调整系数的时频分辨率,并将变换后的系数输出到所述时频滤波系数组织模块;

所述时频滤波系数组织模块,用于将滤波输出的系数按时频平面进行组织,并输出滤波信号。

21、根据权利要求 19 所述的多分辨率矢量量化的音频编码器,其特征在于,所述多分辨率矢量量化器包括矢量组织模块、矢量选择模块、全局归一化模块、局部归一化模块和量化模块;

所述矢量组织模块,用于将所述多分辨率滤波器输出的时频平面系数根据不同的划分策略组织成矢量的形式,输出到所述矢量选择模块;

所述矢量选择模块,用于根据能量的大小等因素选择出待量化的矢量,输出到所述全局归一化模块;

所述全局归一化模块,用于对上述矢量进行全局归一化处理;

所述局部归一化模块,用于计算每个矢量的局部归一化因子,并对所述全局归一化模块输出的矢量进行局部归一化处理,输出到所述量化模块;

所述量化模块,用于对经过两次归一化后的矢量进行量化,并计算量化后的残差。

22、一种多分辨率矢量量化的音频解码器,其特征在于,包括解码和逆量化器、多分辨率逆矢量量化器、多分辨率逆向滤波器和频率-时间映射器;

所述解码和逆量化器，用于对码流解复用、熵解码和逆量化，得到边信息及编码数据，输出到所述多分辨率逆矢量量化器中；

所述多分辨率逆矢量量化器，用于进行逆矢量量化过程，重构量化的矢量；并且将重构矢量加到时频平面上的残差系数，输出到所述多分辨率逆向滤波器；

5 所述多分辨率逆向滤波器，用于对所述多分辨率矢量量化器重构的矢量进行逆向滤波，并输出到所述频率-时间映射器；

所述频率-时间映射器，用于完成信号从频率到时间的映射，得到最终重构的音频信号。

23、根据权利要求 22 所述的多分辨率矢量量化的音频解码器，其特征在于，所述多分辨率
10 逆矢量量化器包括解复用模块、逆量化模块、归一化矢量计算模块、矢量重构模块和加法模块；

所述解复用模块，用于对接收到的码流进行解复用，获得归一化因子和选择点的量化索引；

15 所述逆量化模块，用于根据所述解复用模块输出的信息获取能量包络、矢量量化位置信息，并进行逆量化获取引导点和选择点矢量，计算出二次归一化因子，输出到所述归一化矢量计算模块；

所述归一化矢量计算模块，用于对选择点矢量进行逆二次归一化，获得归一化矢量，输出到所述矢量重构模块中；

所述矢量重构模块，用于根据能量包络对归一化矢量进行逆一次归一化，获得重构矢量；

20 所述加法模块，用于将所述矢量重构模块输出的重构矢量与对应时频平面的反量化残差相加，得到逆量化的时频系数，作为所述多分辨率逆向滤波器的输入。

24、根据权利要求 22 所述的多分辨率矢量量化的音频解码器，其特征在于，所述多分辨率
25 逆向滤波器进一步包括：时频系数组织模块、N 个多分辨率综合模块和 M 个等带宽余弦调制滤波器，且满足 $M=N+1$ ；

所述时频系数组织模块，用于将逆量化系数按滤波输入方式进行组织，如果是缓变信号，则输出到所述等带宽余弦调制滤波器；如果是快变信号，则输出到所述多分辨率综合模块；

所述多分辨率综合模块，用于将多分辨率时频系数映射成等带宽的余弦调制滤波系数，并输出到所述等带宽余弦调制滤波器；

30 所述等带宽余弦调制滤波器，用于对信号进行滤波，获得时域脉冲编码调制输出。

1/8

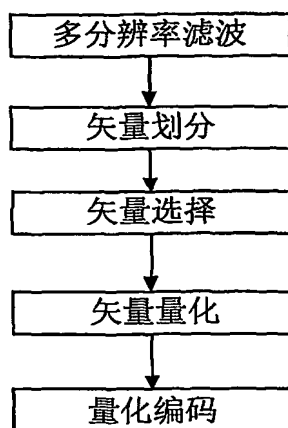


图 1

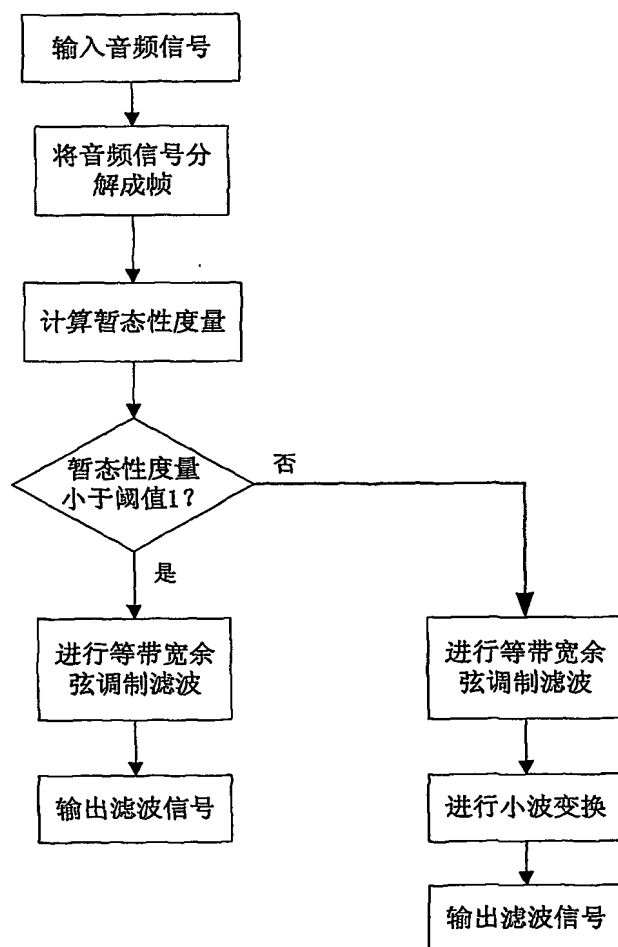


图 2

2/8

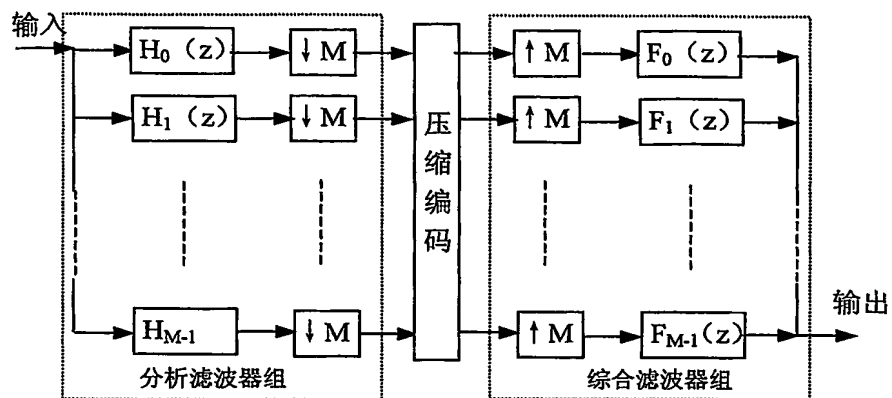
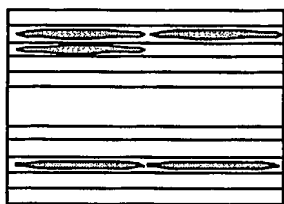
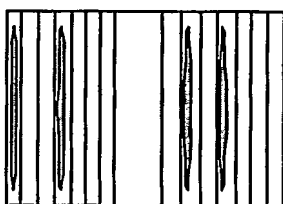


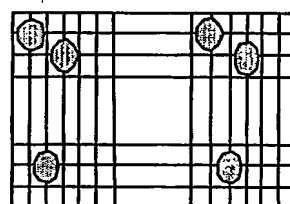
图 3



a



b



c

图 4

3/8

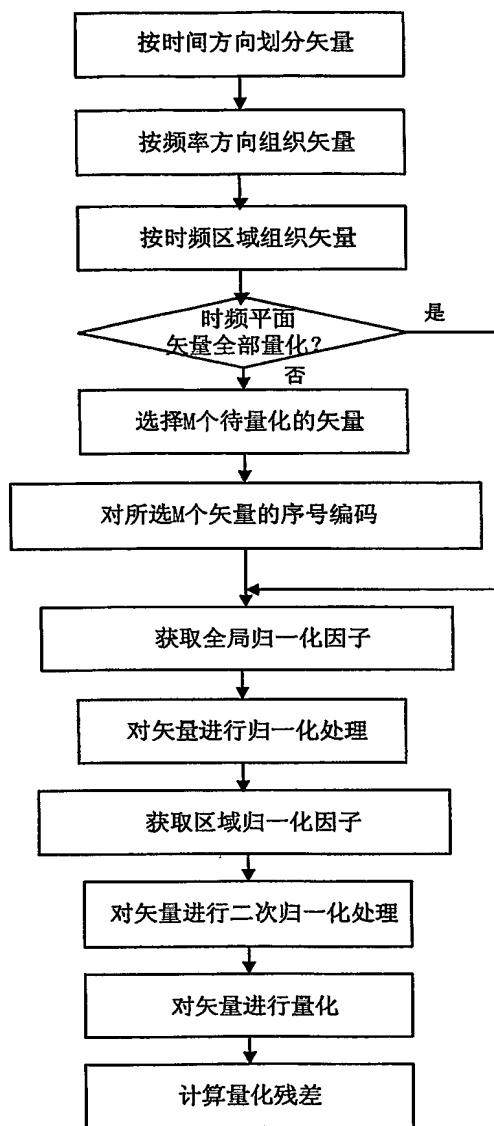


图 5

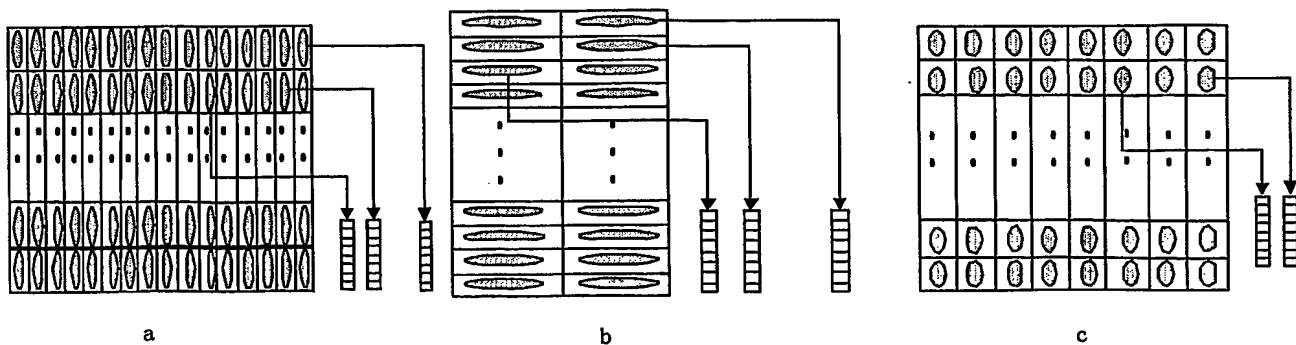


图 6

4/8

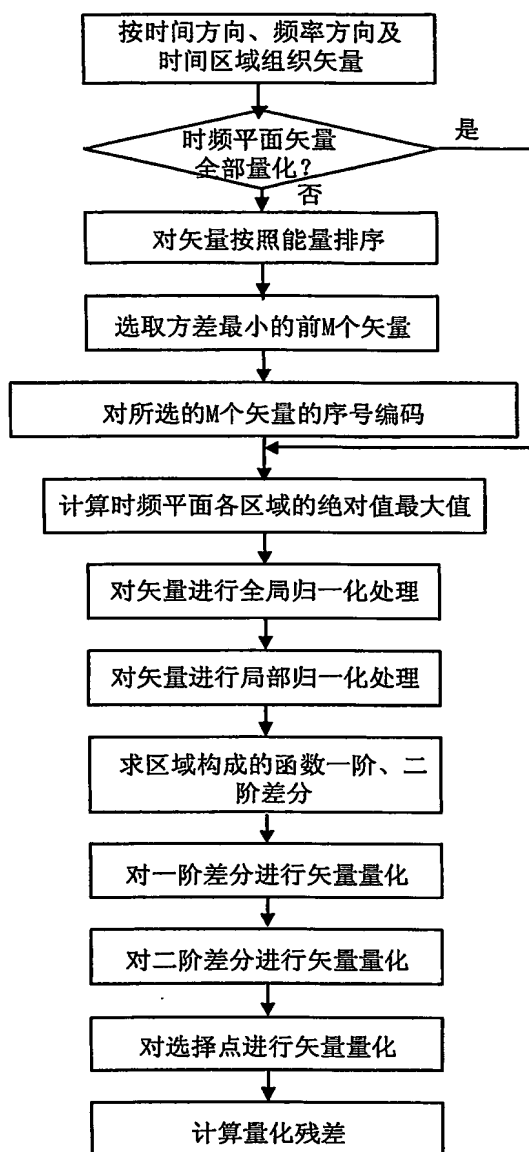


图 7

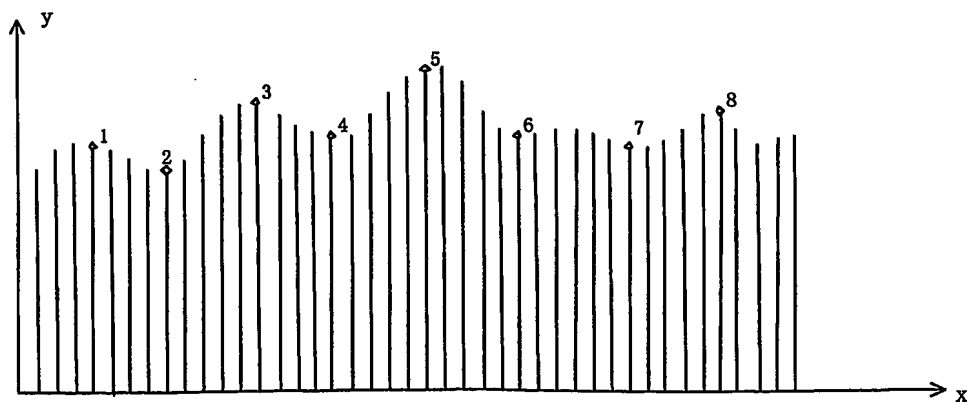


图 8

5/8

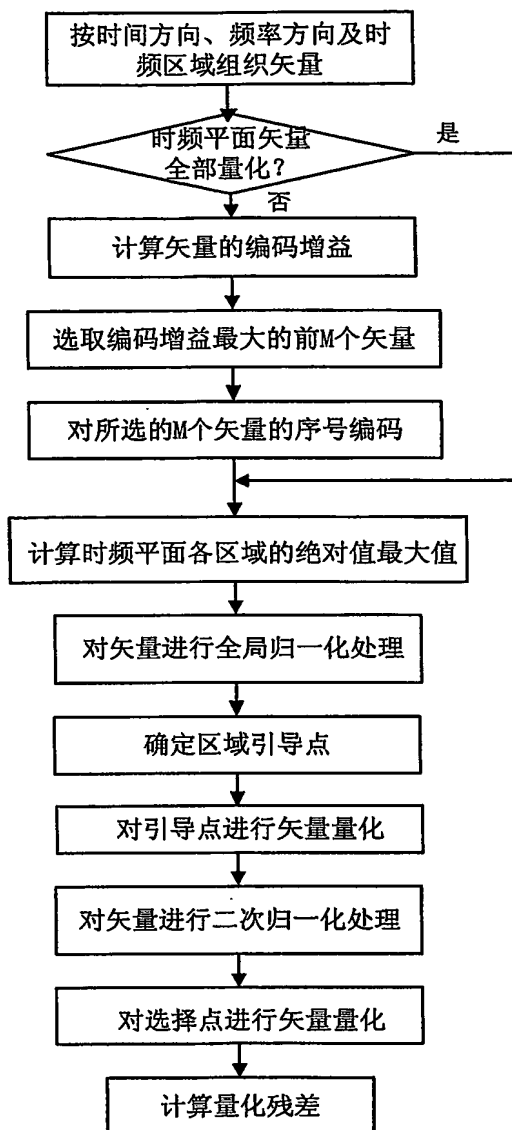


图 9

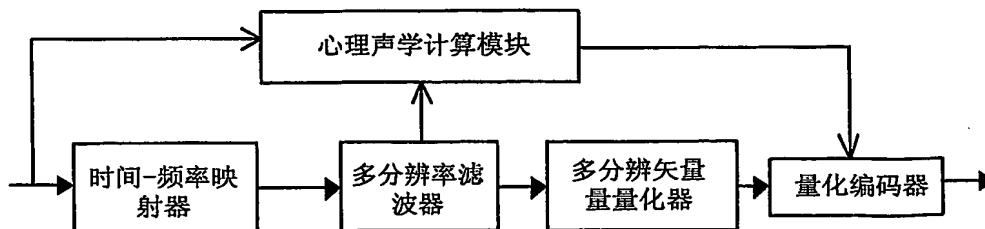


图 10

6/8

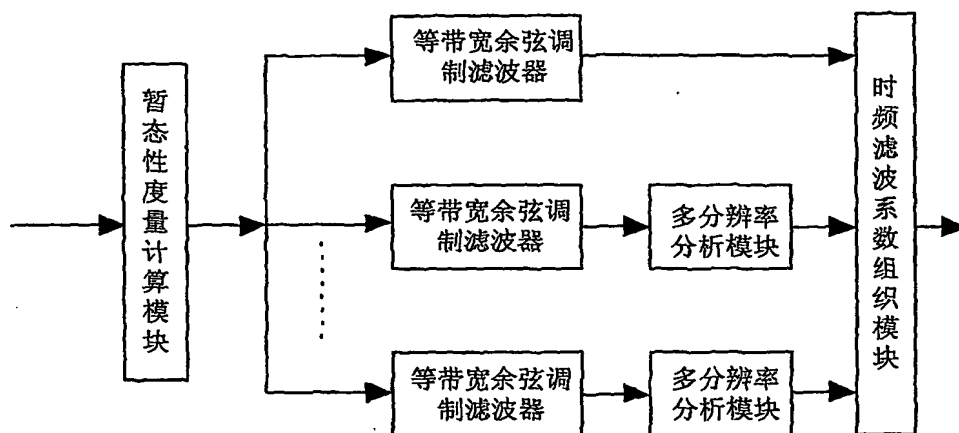


图 11

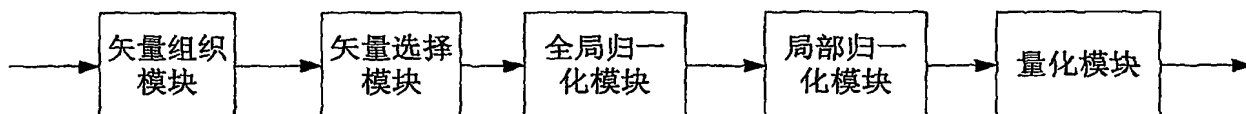


图 12

7/8

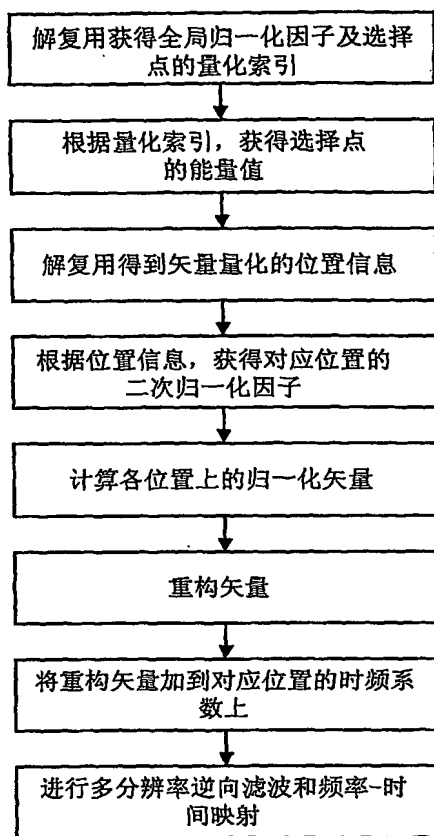


图 13

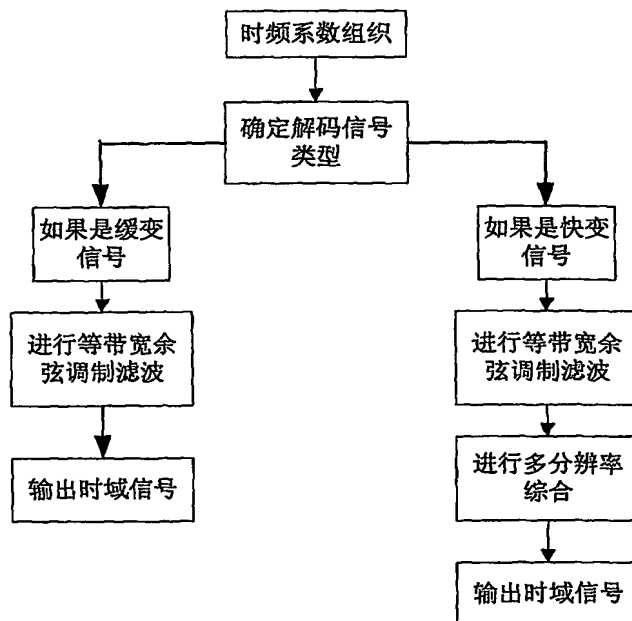


图 14

8/8

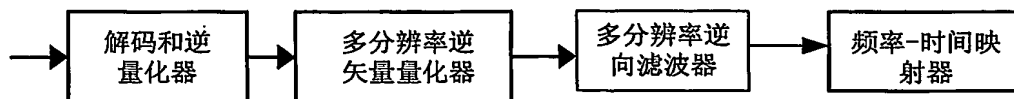


图 15

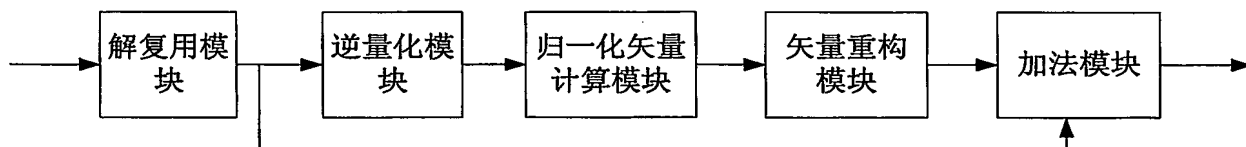


图 16

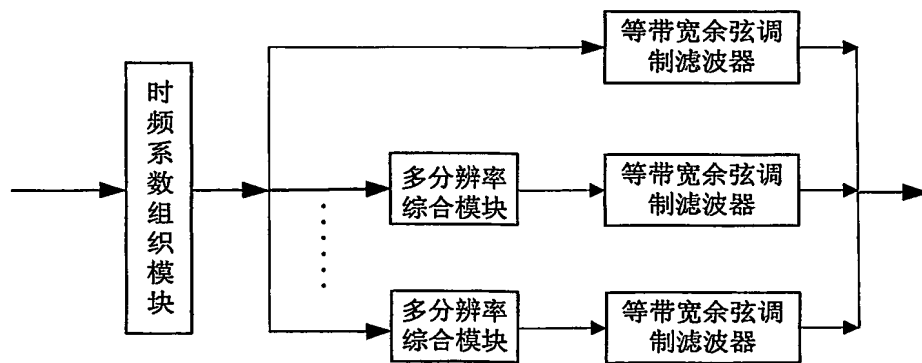


图 17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN03/00790

A. CLASSIFICATION OF SUBJECT MATTER

IPC⁷: G10L19/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC⁷: G10L19/00—19/14

IPC⁶: G10L 7/00-7/06 9/00—9/18

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

CNPAT(1985-)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPODOC WPI PAJ audio code encode decode multiresolution vector quantilization VQ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN,A,1224523 (MATSUSHITA DENKI SANGYO KK) 28.JUL.1999 (28.07.1999),SEE FULL TEXT	1-24
A	CN,A,1222997 (MATSUSHITA DENKI SANGYO KK) 14 .JUL.1999 (14.07.1999),SEE FULL TEXT	1-24
A	US,A,5473727 (SONY CORP) 5.DEC.1995 (05.12.1995) , SEE FULL TEXT	1-24
A	ELECTRONIC AUDIO TECHNOLOGY, 2003.2, PAN XINGDE ZHU XIAOMING "EAC audio coding technology" PAGE 11—15.	1-24

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
01.JUN.2004 (01.06.04)

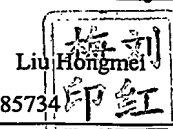
Date of mailing of the international search report

24 · JUN 2004 (24 · 06 · 2004)

Name and mailing address of the ISA/
6 Xitucheng Rd., Jimen Bridge, Haidian District,
100088 Beijing, China

Facsimile No. 86-10-62085303

Authorized officer



Telephone No. 86-10-62085734

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN03/00790

CN 1224523 A	28-07-1999	KR 309727 B	29-09-2001
		WO 9852188 A1	19-11-1998
		JP 11030998 A	02-02-1999
		EP 0919989 A1	02-06-1999
		MX 9810783 A1	01-01-1999
CN 1222997 A	14-07-1999	EP 0910067 B1	13-08-2003
		WO 9800837 A1	08-01-1998
		JP 10020898 A	23-01-1998
		KR 283547 B	02-04-2001
US 5473727 A	05-12-1995	JP 6149296 A	27-05-1994
		JP 3343965 B2	11-11-2002

国际检索报告

国际申请号

PCT/CN03/00790

A. 主题的分类

IPC⁷: G10L19/00

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC⁷: G10L19/00—19/14IPC⁶: G10L 7/00-7/06 9/00—9/18

包含在检索领域中的除最低限度文献以外的检索文献

中国专利文献(1985—)

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

EPODOC WPI PAJ audio code encode decode multiresolution vector quantilization VQ

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN, A, 1224523 (松下电器产业株式会社) 1999 年 7 月 28 日 (28.7.99) 全文	1—24
A	CN, A, 1222997 (松下电器产业株式会社) 1999 年 7 月 14 日 (14.7.99) 全文	1—24
A	US, A, 5473727 (SONY 公司) 1995 年 12 月 5 日 (05.12.95) 全文	1—24
A	《电声技术》2003 年第 02 期, 潘兴德, 朱晓明 “EAC 音频编码技术”, 第 11—15 页。	1—24

☐ 其余文件在 C 栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“B” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

01. 6 月 2004

国际检索报告邮寄日期

24 · 6 月 2004 (24 · 06 · 2004)

中华人民共和国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

授权官员

梅刘
刘红梅

电话号码: (86-10) 62085734

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN03/00790

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN 1224523 A	28-07-1999	KR 309727 B	29-09-2001
		WO 9852188 A1	19-11-1998
		JP 11030998 A	02-02-1999
		EP 0919989 A1	02-06-1999
		MX 9810783 A1	01-01-1999
CN 1222997 A	14-07-1999	EP 0910067 B1	13-08-2003
		WO9800837 A1	08-01-1998
		JP 10020898 A	23-01-1998
		KR 283547 B	02-04-2001
US 5473727 A	05-12-1995	JP 6149296 A	27-05-1994
		JP 3343965 B2	11-11-2002

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☒ **FADED TEXT OR DRAWING**

☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.